# Estimating Causal Effects by Bounding Confounding – Supplementary Material –

## 1 Proof of Lemma 1

**Lemma 1.** *Under the assumption made above, the joint distribution of $X_1, X_2, X_3$ induced by a causal model $M$ or any post-interventional model $M_{\text{do}\, X_i=x}$ has a density w.r.t. the Lebesgue measure (in the continuous case) or counting measure (in the discrete case), respectively. Moreover, this density factorizes according to the causal DAG belonging to the respective model.*

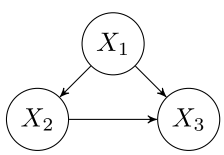


Figure 1: W.l.o.g. we assume this causal DAG.

*Proof.* We only treat the continuous case, the discrete case is straight forward.

Recall our assumption that given a causal model $M$ with causal DAG $G = (V, E)$, for each $X_i \in V$, the random variable $f_i(\text{pa}_i, N_i)$ has a density $q_i(x_i; \text{pa}_i)$ w.r.t. the Lebesgue measure. Note that this implies, that also in any post-interventional model $M_{\text{do}\, X_i=x}$, the random variables $f_i^{M_{\text{do}\, X_i=x}}(\text{pa}_i^{M_{\text{do}\, X_i=x}}, N_i)$ have densities w.r.t. the Lebesgue measure which can easily obtaine from the $q_i(x_i; \text{pa}_i)$. Hence, w.l.o.g., we only prove the lemma w.r.t. $M$.

In what follows, we will only consider the case where the causal DAG is fully connected, the other cases work similarly. W.l.o.g. we assume the DAG in Figure 1.

Let $q(x_1, x_2, x_3) := \prod_i q_i(x_i; \text{pa}_i)$.

To see that $q(x_1, x_2, x_3)$ factorizes according to $G$, note that

$$\begin{aligned} p(x_3|x_2, x_1) &= \frac{q_3(x_3; x_1, x_2) q_2(x_2; x_1) q_1(x_1)}{\int q_3(x_3; x_1, x_2) q_2(x_2; x_1) q_1(x_1) \mathrm{d}x_3} \\ &= \frac{q_3(x_3; x_1, x_2) q_2(x_2; x_1) q_1(x_1)}{q_2(x_2; x_1) q_1(x_1)} \\ &= q_3(x_3; x_1, x_2). \end{aligned}$$

Similary one calculates $p(x_2|x_1) = q(x_2; x_1)$ and $p(x_1) = q(x_1)$.

It remains to show that $q(x_1, x_2, x_3)$ it is a density for the joint distribution $P(X_1, X_2, X_3)$.

Keep in mind that for measurable $f, Y$ we have [Bogachev, 2007]

$$\int Y(s) \mathrm{d}P_{f(N)}(s) = \int Y(f(r)) \mathrm{d}P_N(r). \qquad (1)$$

Let $[\cdot]$ denote the characteristic function (i.e. it equals 1 if the statement inside the brackets is true and 0 otherwise). Now we can calculate

$$
\begin{aligned}
&\int_{-\infty}^{a}\int_{-\infty}^{b}\int_{-\infty}^{c} q_1(x_1)q_2(x_2;x_1)q_3(x_3;x_1,x_2)\mathrm{d}x_3\mathrm{d}x_2\mathrm{d}x_1 && (2)\\
&=\int[x_1\le a]\int[x_2\le b]\int[x_3\le c]\mathrm{d}P_{f_3(x_1,x_2,N_3)}(x_3)\mathrm{d}P_{f_2(x_1,N_2)}(x_2)\mathrm{d}P_{f_1(N_1)}(x_1) && (3)\\
&=\int[x_1\le a]\int[x_2\le b]\int[f_3(x_1,x_2,n_3)\le c]\mathrm{d}P_{N_3}(n_3)\mathrm{d}P_{f_2(x_1,N_2)}(x_2)\mathrm{d}P_{f_1(N_1)}(x_1) && (4)\\
&=\int[x_1\le a]\int[f_2(x_1,n_2)\le b]\int[f_3(x_1,f_2(x_1,n_2),n_3)\le c]\mathrm{d}P_{N_3}(n_3)\mathrm{d}P_{N_2}(n_2)\mathrm{d}P_{f_1(N_1)}(x_1) && (5)\\
&=\int[f_1(n_1)\le a]\int[f_2(f_1(n_1),n_2)\le b]\int[f_3(f_1(n_1),f_2(f_1(n_1),n_2),n_3)\le c]\mathrm{d}P_{N_3}(n_3)\mathrm{d}P_{N_2}(n_2)\mathrm{d}P_{N_1}(n_1) && (6)\\
&=\int[f_1(n_1)\le a][f_2(f_1(n_1),n_2)\le b][f_3(f_1(n_1),f_2(f_1(n_1),n_2),n_3)\le c]\mathrm{d}P_{N_1,N_2,N_3}(n_1,n_2,n_3) && (7)\\
&=\mathbb{E}\big[[f_1(N_1)\le a][f_2(f_1(N_1),N_2)\le b][f_3(f_1(N_1),f_2(f_1(N_1),N_2),N_3)\le c]\big] && (8)\\
&=P(X_1\le a,X_2\le b,X_3\le c), && (9)
\end{aligned}
$$

where equations (4), (5), (6) follow by applying equation (1), and equation (7) follow from the independence of the noise terms $N_i$.

This proofs that $q(x_1,x_2,x_3)$ is a density of $P(X_1,X_2,X_3)$ w.r.t. the Lebesgue measure.

□

## 2 Proof of Lemma 2

**Lemma 2.** *For all $x$ we have*

$$
\begin{aligned}
p(Y|X=x,\mathrm{do}\,X{=}x) &= p(Y|X=x),\\
\mathbb{E}[Y|X=x,\mathrm{do}\,X{=}x] &= \mathbb{E}[Y|X=x].
\end{aligned}
$$

*Proof.* Based on the proof for Lemma 1, we have

$$
\begin{aligned}
p(u,x',y|\mathrm{do}\,X{=}x') &= q_U(u)q_X(x';u)q_Y(y;u,x') && (10)\\
&= p(u)p(x'|u)p(y|u,x'), && (11)
\end{aligned}
$$

where equation (10) holds true because $q_U(u)$, $q_X(x;u)$ and $q_Y(y;u,x')$ are the densities for $f_U^{M_{\mathrm{do}\,X_i=x'}}(N_U)$, $f_X^{M_{\mathrm{do}\,X_i=x'}}(u,N_X)$ and $f_Y^{M_{\mathrm{do}\,X_i=x'}}(u,N_Y)$, respectively. Equation (11) implies that

$$p(u,x',y|\mathrm{do}\,X{=}x') = p(u,x',y),$$

and hence

$$p(y|X=x',\mathrm{do}\,X{=}x') = p(y|x').$$

□

## 3 Proof of Theorem 5

**Theorem 5.** *For all $x$*

$$\sqrt{\mathcal{F}_{Y|X}(x)} - \sqrt{\mathcal{F}^2_{Y|X,\mathrm{do}\,X}(x,x)} \le \sqrt{\mathcal{F}^1_{Y|X,\mathrm{do}\,X}(x,x)}.$$

*Proof.* First note that by the chain rule

$$
\begin{aligned}
&\mathrm{d}_x \log p(y|X{=}x,\mathrm{do}\,X{=}x)\\
&= \partial_1 \log p(y|X{=}x,\mathrm{do}\,X{=}x)\\
&\quad + \partial_2 \log p(y|X{=}x,\mathrm{do}\,X{=}x).
\end{aligned}
$$

By Lemma 2 we have $p(y|X{=}x) = p(y|X{=}x,\mathrm{do}\,X{=}x)$ for all $x, y$.

Together we obtain

$$
\begin{aligned}
&\left(\mathbb{E}[(\mathrm{d}_x \log p(y|X{=}x))^2]\right)^{\frac{1}{2}}\\
&= (\mathbb{E}[(\partial_1 p(y|X{=}x,\mathrm{do}\,X{=}x)\\
&\quad +\partial_2 p(y|X{=}x,\mathrm{do}\,X{=}x))^2]\big)^{\frac{1}{2}}\\
&\le \left(\mathbb{E}[(\partial_1 p(y|X{=}x,\mathrm{do}\,X{=}x))^2]\right)^{\frac{1}{2}}\\
&\quad + \left(\mathbb{E}[(\partial_2 p(y|X{=}x,\mathrm{do}\,X{=}x))^2]\right)^{\frac{1}{2}}.
\end{aligned}
$$

Note that the expectation is taken w.r.t. $p(y|x)$. □

## 4 Proof of Proposition 1

**Proposition 1.** *In the given scenario we have* $\mathrm{I}(U:X) \le \mathrm{H}(X)p(W{=}1)$.

*Proof.* We calculate

$$\begin{aligned}
&\mathrm{I}(U : X) \leq \mathrm{I}(U : X) + \mathrm{I}(U : W|X) = \mathrm{I}(U : W, X) \\
&= \mathrm{I}(U : W) + \mathrm{I}(U : X|W) \\
&= \mathrm{I}(U : X|W{=}0)p(W{=}0) + \mathrm{I}(U : X|W{=}1)p(W{=}1) \\
&= \mathrm{I}(U : X|W{=}1)p(W{=}1) \leq \mathrm{H}(X)p(W{=}1).
\end{aligned}$$

□

## References

V. Bogachev. *Measure Theory*. Springer, 2007.